Reply to reviewer 2:

Dear Reviewer,

we would like to thank you for taking the time to carefully review our manuscript. Below you will find our response to your comments in blue indicating the respective changes/improvements where necessary.

Spiegl et al. examine the impact of the 11-year solar cycle on decadal predictions, by using a large data-set from the MPI decadal prediction system. The length and size of their ensemble experiments offer an unprecedented opportunity to critically evaluate the possibility (suggested in many other studies) concerning a possible role of the solar cycle as a source of predictability (via the "top-down" mechanism). This is a very controversial and yet relevant topic, given the potential implicatuions for decadal prediction of boreal wintertime climate in the North Atlantic. This paper nicely shows that the top-down signal in the MPI model is small and there is hardly any synchronization of the NAO. In particular, the paper also debunks the possibility raised in another recent paper (Drews et al., 2022) that the time-dependence of the solar/NAO signal highlighted in Chiodo et al. (2019) might only arise in decades with a stronger solar cycle. The paper nicely shows that indeed, the modeled response is weak and mostly not significant and indeed, the main conclusion against a solar modulation of the NAO and synchronization for their model is absolutely warranted. I think that the paper will contribute to the scientific debate on this matter and deserves to be published, after some corrections as detailed below.

GENERAL COMMENTS

- The main conclusions of the paper are supported by the evidence provided in the paper. However, it would be good to at least acnowledge and/or discuss the possibility that the MPI-ESM-HR model system, just like other model systems, might suffer from the "signal-to-noise" paradox issue, according to which, the models might underestimate any externally forced signals in dynamical aspects of the climate system (see e.g., Scaife and Smith, https://www.nature.com/articles/s41612-018-0038-4), such as the PNJ and stratosphere-troposphere coupling.

Thanks for this important point! Added the following paragraph to the discussion section:

Added: "The discrepancies between the observed and modelled internal variability in response to external forcings (such as solar variability) may also be attributed to the "signal-to-noise" paradox which states that relatively small changes in the external forcing will not lead to detectable changes in the variability spectrum in both the real climate system and model simulations as discussed by Scaife and Smith (2018)."

- It might be also good to at least discuss the implications of the part of the solar forcing which is missing in these runs; the particles forcing (MEE/SEP). I would argue that particles are unlikely to change the picture and the EPP signal is perhaps even more complicated than the irradiance/UV signal, and the models are at their infancy in simulating this forcing... but it needs to be at least discussed, given the number of related publications arguing that EPP play an important role in the solar influences on the polar vortex/ NAM and NAO.

The effects of solar energetic particles (SEP) and medium energy electrons (MEE) were not explicitly included in the model, as in our opinion – as also mentioned by the reviewer - these effects are unlikely to change our results, since they probably are (if present) even smaller and more complex than the alleged 11-year solar cycle surface signals. On the other hand, the model uses the observed ozone time series as a boundary condition, which might have been influenced by SEP and MEE.

To complete the discussion we added:

Added:" While the model simulations include both, changes in the total solar irradiance (TSI) and spectral solar irradiance (SSI), potential effects related to solar energetic particles (SEP) and medium energy electrons (MEE) are not explicitly included in the MiKlip experiments. Observations and model studies suggest that changes in the stratospheric composition related to SEP can lead to a radiatively driven modulation of the middle atmosphere dynamics, which can penetrate to lower atmospheric layers down to the troposphere (e.g., Seppälä et al., 2009, 2014; Baumgaertner et al., 2010; Arsenovic et al., 2016). However, since no robust surface impacts have been simulated even for strong solar energetic particle events (SEP) of the recent decades (Jackman et al., 2009), we infer that including these effects may not alter our results significantly.

The other general point concerns the lengthy discussions of the individual months. I think many of the parts of sections 5-6 could be shortened, especially if there's no robust signal to document.

We prefer to not shorten Sections 5 and 6 for the sake of completeness and to avoid misinterpretation given the complexity of the subject and the ongoing discussions in the community.

-Lastly, the paper examines the top-down mechanism in detail and rules out a robust effect in their model... but what about any bottom-up impacts? They should be at least mentioned somewhere in the discussion, to put this paper in the broader context. The existing literature is cited but there could be more discussion of the comparison with Chiodo et al., 2019 and the aspects that are new here (i.e. testing the link between the time-dependence of the signal and the magnitude of the solar cycle amplitude).

Indeed, the paper mainly discusses possible top-down induced solar signals. However, since the model also includes TSI variations (the main driver for eventual bottom-up mechanisms), bottom-up effects would be included if they shape the decadal "surface solar signal". However, we can't find any systematic pattern in the surface meteorological variables nor the NAO. Additionally, to our knowledge, there is no literature available that addresses potential bottom-up effects in the North Atlantic Sector, since profound TSI effects are mainly limited to the low latitudes. We now mention bottom-up effects in the discussion section and use the argumentation as above. Additionally, we explicitly compare the strategy of our study to Chiodo et al., (2019) and include the point mentioned by the reviewer above.

Added:

" The MiKlip simulations are more in line with Chiodo et al. (2019), who argued that the alleged surface solar signals could be an incidental product which is only detectable during phases with stronger solar cycles. Our results even suggest that robust solar surface imprints are basically absent throughout the complete historical period and are thus not sensitive to the amplitude of individual solar cycles."

and

"It should be noted, that we did not explicitly analyse a potential TSI controlled bottom-up effect on the solar surface signal, as bottom-up effects are rather confined to tropical latitudes with a prolonged influence of the TSI throughout the year (Meehl et al., 2008)."

SPECIFIC COMMENTS

L17-19: could add that previous work didn't really do a decadal prediction set-up --> rather, continuous long-term climate runs!

We agree. A short note has been added to the abstract.

L20 - it's more than just "confirmation" from other modeling groups - a cleaner model study is really missing...!

Adjusted.

L22 we aim for an unbiased evaluation --> "unbiased" seems a bit too harsh and indirect criticism towards previous studies --> replace with "objective and improved"...?

Done.

L30 is rather weak -> suggest removing "rather"

Done.

L31 remove "basically"

Done.

L36-37 unclear what the driver of the "anomalies" is - a bit more clarity would be needed. Would these be ensemble mean anomalies (or anomalies in the individual ensembles) that are correlated with the solar cycle, or what is the "driver" of these anomalies..?

Rephrased this paragraph for more clarity.

Rephrased:" We find that the westerly wind anomalies in the lower troposphere as well as the anomalies in the mean sea level pressure are most likely independent from the timing of the seasonal march in the middle atmosphere and thus alleged top-down influences. The pattern rather reflects the decadal internal variability of the troposphere, mimicking positive and negative phases of the Arctic- and North Atlantic Oscillation throughout the year sporadically, which are then assigned to the solar predictor time series without any physical plausible connection and sound solar contribution."

L36 "most likely independent from the seasonal march in the middle atmosphere" -> would this mean that the timing of the tropospheric anomalies does not match the downward propagation of the (apparent) solar signals? If so, I'd reword this to use the word "timing" so that it's clearer to the reader what we are talking about.

See above.

L38 "rather sporadically than in a systematic way" -> sporadic in that they depend on the ensemble member, or sporadic in that they depend on the time window being analysed? again, more specificity might be appreciated.

See above.

L40 "might rather be interpreted" - I'd repharase to "might be". Just remove "interpreted" to simplify the wording of this part of the abstract.

Done.

L41 "between the solar forcing" - it might be good to again be more specific and say that the paper handles the UV/irradiance forcing specifically... as there might be also other pathways for a solar influence on climate, namely via MEE/particle forcing. Just add "irradiance" or "UV" to "solar forcing"

Done.

L40 "as a statistical artefact" - it may well be that the synchronization is simply due to internal dynamics of the atmosphere...? if that's the case, then it may not be a "statistical artifact" in the strict sense, but rather, that it's not solar driven...

Yes, that's what we mean.

The wording has changed to: "statistical artefact, affected for example by the internal decadal variability of the ocean… "

L73 "...convergence in the Eliassen-Palm flux (EPF)...positive wind anomaly" a convergence in EPF would result in a deceleration of U. I guess what is meant here is thus the "divergence" of the EPF?

Yes, correct. Thank you for spotting this.

L77 - Discussion of Matthes 2004 - I'd be more specific here and say that this study specifically studied the evolution of the signal on sub-monthly time-scales

Matthes et al. (2004) analyzed monthly means, while Matthes et al. (2006) used monthly subperiods to derive the downward transfer of the solar signal into the troposphere. A short note has been added.

L81 "from February on" -> from February onwards

Done.

L83 "very individual temporal progressions" - again, I'd use the word "timing" to clarify what is meant here. E.g. you could reword this to "the exact timing of the downward propagation depends on the individual study"

The sentence has been changed as suggested.

Side note: I'm not sure Marsh '07 really show a downward solar impact - rather, they show, by means of a time-slice experiment, that the signal is robust in the upper stratosphere but in the polar regions, their signal is rather weak and non-significant...

The reviewer is correct! What Marsh et al. (2007) are showing is the long-term annual mean response to the solar cycle. Based on this, no conclusions can be drawn with respect to the "top-down mechanism". Thus, this citation has been removed.

around L90: it might be good to add somewhere here that there are also other potential pathways for a solar influence on climate but the irradiance-UV component is the most studied one and the one that models have an easier time simulating...?

We agree that there are other pathways of the solar signal. This paper, however, is about the top-down mechanism and the North Atlantic sector. Discussing e.g., bottom-up effects which "might" be relevant in tropics seem misplaced in this context.

L110 I think you could also add here that the availability of the large ensembles with the observed time-varying solar cycles allows you to explicitly test the dependence of the signal on the solar cycle amplitude (hypothesis put forward by Drews et al 2022).

We have added this, thanks.

L209 I'd suggest removing unscientific terms such as "pretty".

The sentence has been removed and replaced.

L219 replace "receive" with "obtain"

Done.

Figure 2 seems a bit blurry and the individual components of the figure (e.g. colorbar and right panel) seem to have been manually assembled together. I'd suggest including a high resolution version of this figure and use another software to produce the figure without having to manually assemble the individual panels. This also concerns the other Figs. 3-8.

The reviewer is correct! Figure 2 has not been ideal yet! Replaced by another version. With respect to the other figures, we included the grey bars manually and we would like to keep this style. We slightly adjusted the arrangement of the subplots though. If the individual plots seem not to be high resolution figures in the manuscript yet, this would be due to the fact that the figures have been copied/paced from .eps and then converted to .pdf. The final figures that will be uploaded separately at the end will of course be in higher resolution to meet the quality standards of WCD.

Figure 3 nice figure, but I have the impression that the data is not uniformly scattered in the x-dimension (Sunspot number) - the dots are all structured in "vertical bands". Why is that happening? Is the SSN sampled only every 5-10 units...? Also, do dots correspond to each individual member?

There is a misunderstanding. The dots do not only correspond to individual ensemble members but represent all model years. The data are also not only sampled only every 5-10 units. The "vertical bands" appear because the same SSN time series (for each ensemble member) has been used for correlation. The response, however, is individual in each ensemble member. This can be seen best during the SC19 maximum, which is characterized by the highest SSN number throughout the historical epoch. This SSN value appears exactly once in 120 years (lower values appear more often), thus exactly 10 dots are arranged around this value (sometimes stronger, sometimes weaker responses). The x-axis would be correct though.

L410-440 (and intro) I think it would also be good to add the most recent paper by Gray 2016 to this discussion... and in particular, Ma et al., 2021 - who argued that the early-winter signal is the most robust component of the signal.

Since Gray et al., 2016 is primarily about blocking frequencies, we added Ma et al. (2018) and added it to the discussion of the early winter influence. We could not find a paper of Ma et al. (2021) dealing with solar variability and the North Atlantic sector.

Added:" A most recent study again concludes that the most pronounced solar signal seem to appear in early winter (Ma et al. (2018))."

L441 I think you can be more confident here and replace "assume" with "conclude"

Done.

L444-447 What about the "bottom up" signal? I agree that much of this discussions disqualifies the top-down... but can we also rule out a role of any direct surface (TSI-driven) signals?

Please see in the main comments above and thereafter.

L544-546 Could the authors provide some evidence to support the interpretation that the background state determines if the signal is transferred or not?

In our opinion (and at this stage) the very different “solar signals” in the middle atmosphere that we observe in individual ensemble members, even though they have been driven with the exact same external parameters, could already be a hint to the importance of the middle atmosphere dynamics. A more detailed investigation is currently part of the SOLCHECK project. The following sentence has been added: ”The important role of middle atmosphere dynamics in modulating potential solar signals is currently investigated as part of the SOLCHECK project and will be published in a subsequent paper (Wenjuan Huo, personal communication).”

L550 remove "time series"

Done.

L551 not sure what is meant with "manifold phase relations". can the authors clarify?

We mean different NAO-solar forcing phase relations, as shown in Figure 8. The sentence has been clarified: “We find a range of phase relations between the NAO and the solar forcing throughout our ensemble members, which implies a random statistical relation rather than a physical sound connection.”

L563 "...made in model development.." agreed that having larger ensembles helps diagnosing robust signals and probably that's partly the reason why signals in older studies were deemed as "solar"... on the other hand, I do not think that model physics being simpler in those early studies is another reason for the apparently stronger signal. Those models had all the necessary physics to capture the mechanism, in principle.

We fully agree. The early GCMs included the necessary UV-radiation codes and middle atmosphere dynamics to simulate the solar top-down signal. However, to study the tropospheric signal, more components had to be added to the models, in particular ocean models. This together with ensemble realizations nearly inhibits the identification of a robust surface solar signal.

We extended the sentence to: “While these models disposed of the necessary physical mechanisms, i.e., UV radiation codes and middle atmosphere dynamics, to capture the solar UV-induced top-down solar signal, the complex nature of physical and chemical processes and the spectrum of internal variability were reduced.”

L583 I'd add somewhere here that these results agree with the conclusions of Chiodo et al., 2019 - and here, they even strengthen those conclusions concerning the little role of solar forcing in modularing the NAO, as the runs of this paper can even more directly be compared with observations and the solar signal can be tests in an even more realistic set-up (with transient forcings, transient solar cycle, etc.)

Please see our comment, related to the now included comparison to Chiodo et al. (2019) above. Additionally, we added the important point mentioned by the reviewer that the MiKlip simulations are more realistically, since they include all observed forcings over the historical period.